JOURNAL OF BACTERIOLOGY, Jan. 1995, p. 401–412 Vol. 177, No. 2
0021-9193/95/$04.00+0

Discontinuities in the Evolution of *Pseudomonas putida cat* Genes†

JOHN E. HOUGHTON,[1,2] TIMOTHY M. BROWN,[1] ANDREW J. APPEL,[1] E. JOHN HUGHES,[2]‡ AND L. NICHOLAS ORNSTON[2]*

Department of Biology, Georgia State University, Atlanta, Georgia 30303,[1] Department of Biology, Yale University, New Haven, Connecticut 06511[2]

Received 29 July 1994/Accepted 4 November 1994

The organization and transcriptional control of chromosomal *cat* genes (required for dissimilation of catechol by the β-ketoadipate pathway) in the *Pseudomonas putida* biotype strain (ATCC 12633) are reported. Nucleotide sequence reveals that *catR* is separated by 135 bp from the divergently transcribed *catBC,A*; *catC* begins 21 nucleotides downstream from *catB*, and *catA* begins 41 nucleotides downstream from *catC*. This contrasts with the gene arrangement in other bacteria, in which *catA* lies several kilobases upstream from *catB*. Properties of Tn*5* mutants confirmed earlier suggestions that *catR* is a transcriptional activator and indicated that *catA* is activated by CatR independently of its activation of *catBC*. CatR binds to both a DNA fragment containing the *catR-catB* intergenic region and another DNA fragment containing *catC*. *Pseudomonas* strain RB1 resembles *P. putida* in some respects. Divergence of the two *Pseudomonas* chromosomes was revealed as nucleotide substitution of about 10% after alignment of known portions of *catR,BC,A*. Divergent transcriptional controls are suggested by a cluster of nucleotide sequence modifications in *Pseudomonas* strain RB1 which disrupt a stem-loop structure directly upstream of *catB* in the *P. putida* chromosome. Abrupt divergence of the *catR,BC,A* nucleotide sequences was achieved during evolution by insertion of an 85-bp palindromic genetic element uniquely positioned downstream from *P. putida catR* and counterpoised by insertion of a similar palindromic sequence in the *Pseudomonas* strain RB1 *catB-catC* intergenic region. Properties of the palindromic genetic element suggest that it may serve functions analogous to those of repetitive extragenic palindromic sequences and enteric repetitive intergenic consensus sequences in enteric bacteria.

Many investigations of aromatic catabolism have been conducted with the *Pseudomonas putida* biotype strain (ATCC 12633) and have allowed elucidation of the central reactions of the β-ketoadipate pathway (34). In this metabolic system, three reactions are associated uniquely with catechol metabolism (Fig. 1): catechol 1,2-dioxygenase (EC 1.13.1.1) (encoded by *catA*), muconate cycloisomerase (EC 5.5.1.1) (encoded by catB), and muconolactone isomerase (EC 5.3.3.4) (encoded by *catC*). The *P. putida catB* and *catC* nucleotide sequences are of particular interest because they encode proteins for which crystal structures have been determined (14, 17). CatC is unusual in that it contains only 97 amino acids in its primary structure, and the five active sites of the enzyme are formed between pairs of protein subunits within a decameric array (17). The three-dimensional structure of CatB (muconate cycloisomerase) proved to be similar to that of MdlA (mandelate racemase), and comparison of the primary structures of these proteins revealed their common ancestry (28). This finding provided a significant insight into the gene pools that were called upon in metabolic evolution because CatB and MdlA catalyze different reactions in the same metabolic pathway (37).

The *P. putida cat* genes are expressed in response to the metabolite inducer muconate (32) (Fig. 1). Transductional analysis of spontaneous mutants, obtained from the *P. putida* biotype strain derivative PRS2000, showed a clustering of genes in the order *catR*, *catB*, *catC* and indicated that *catR* is required for expression of *catB* and *catC* (54, 56). Physiological studies suggested that *catBC* forms an operon resisting catabolite repression more than *catA* (32). The location of *catA* in *P. putida* PRS2000 was not determined, but several kilobase pairs DNA are known to lie between the independently transcribed *catA* and *catBC* regions of chromosomes from the bacterial species *Pseudomonas aeruginosa* (20) and *Acinetobacter calcoaceticus* (29).

Evidence that would allow for further understanding of biological properties associated with the *P. putida catR,BC* region demanded cloning of the genes. The most direct approach was to select for broad-host-range plasmids carrying a wild-type *P. putida* PRS1 DNA insert that complemented *catR,BC* mutations in the host strain. These efforts, however, were hampered by genetic events that transferred *catR,BC* DNA from donor plasmids into the chromosome, leaving recombinant strains maintaining plasmids in which DNA inserts were no longer detectable. Success was achieved in cloning DNA carrying *catR,BC* (1, 2) from another fluorescent pseudomonad, strain RB1 (57). Comparison of *Pseudomonas* strain RB1 *cat* nucleotide sequences with those of analogous *clc* genes revealed their common ancestry (10). The *clc* genes are plasmid encoded and are responsible for dissimilation of chloroaromatic compounds (10, 11).

Divergence of the *cat* regions within *P. putida* from the corresponding genes in *Pseudomonas* strain RB1 was indicated by the finding that *cat* genes from the latter organism were stably maintained on broad-host-range plasmids after their introduction into mutants derived from the *P. putida* biotype strain (1, 2). To ascertain the extent to which *cat* genes from the *P. putida* biotype strain and from the *Pseudomonas* strain RB1 differ in nucleotide sequence, it was necessary to clone the *cat* region from the biotype strain. As described here, this was achieved by introduction of Tn*5* into the *cat* genes. *P. putida* genes that had been inactivated by Tn*5* insertion were cloned

* Corresponding author. Mailing address: Department of Biology, Yale University, P.O. Box 208103, New Haven, CT 06520-8103. Phone: (203) 432-3498. Fax: (203) 432-6161.

† Publication no. 9 from the Biological Transformation Center in the Yale Biospherics Institute.

‡ Present address: Davies, Collison, and Cave, Melbourne, Victoria 3000, Australia.

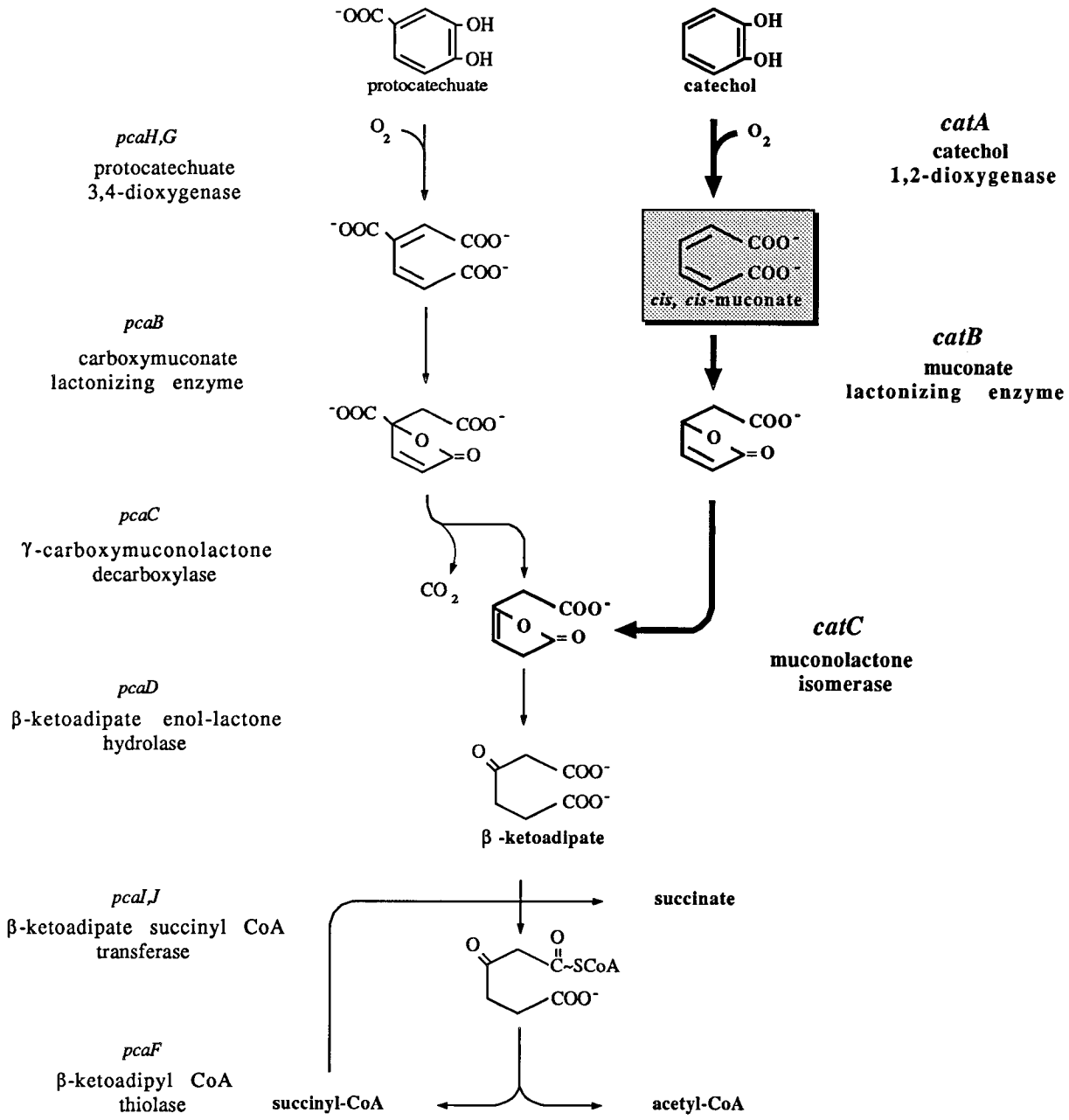

FIG. 1. Enzymatic reactions associated with the *cat* structural genes (represented by heavy arrows) in the metabolism of catechol. In *P. putida*, expression of the genes is elicited by the metabolite *cis,cis*-muconate (boxed) under control of the transcriptional activator *catR*. Metabolism of catechol converges with the pathway for utilization of protocatechuate by enzymes (encoded by the *pca* structural genes). Primary growth substrates used in this investigation include benzoate, which is metabolized via catechol, and *p*-hydroxybenzoate, which is metabolized via protocatechuate. CoA, coenzyme A.

in *Escherichia coli* and subsequently used as probes to identify other clones carrying unaltered, wild-type *cat* DNA. Characterization of the Tn*5* mutant strains of *P. putida* fortifies the earlier conclusion that *catR* functions as a transcriptional activator. In addition, examination of the cloned DNA reveals an unpredicted gene organization: *catA*, which is subject to transcriptional controls independent of those exerted over *catBC*, lies directly downstream from *catC*. Nucleotide sequence comparisons reveal regions in which divergence of the *cat* genes appears to have been abrupt.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1. All of the *P. putida* strains were derived from the *P. putida* biotype strain PRS1 through its derivative strain PRS2000 (33). Plasmid pRKTV14 and its host *E. coli* Neco100 were gifts from N. Panopoulos (58).

Media and growth conditions. *E. coli* and *P. putida* cultures were grown aerobically with constant shaking at 37 and 30°C, respectively, in either L broth (27) or defined mineral medium (34). Unless specified otherwise, aromatic compounds and other organic growth substrates were supplied at concentrations of 5 and 10 mM, respectively. When required for selection, kanamycin was added to a final concentration of 100 μg/ml for *P. putida* and 50 μg/ml for *E. coli* strains; ampicillin was used at a final concentration of 50 μg/ml. For enzyme analysis, *P.*

TABLE 1. Bacterial strains and plasmids used

Strain or plasmid	Relevant genotype and phenotype[a]	Reference or source
Strains		
P. putida		
PRS2000	Derived from *P. putida* biotype strain PRS1 (ATCC 12633)	54
PRS2014	*catR1122* (Ben^- Pob^+)	54
PRS2015	Δ*catB1123* (Ben^- Pob^+)	54
PRS3002	*catC102*::Tn*5* Km^r (Ben^- Pob^+)	This study
PRS3003	catB103::Tn*5* Km^r (Ben^- Pob^+)	This study
PRS3026	*catR126*::Tn*5* Km^r (Ben^- Pob^+)	This study
E. coli		
Neco100	*leu thr thi trp recA lacY*	58
JM109	recA endA gyrA96 thi hdsR17 supE44 relA lacY (*lac proAB*) [F′ *traD36 proAB lacI*^{q} *Z* M15]	53
DH5α	F^- p80, Δ*lacZ* DM15 Δ(*lacZYA-argF*) U169 *ind1 recA1 hsdR17* (r_K^-, m_K^+) *deoR thi-1 supE44 gyrA96 relA1*	Bethesda Research Laboratories
BL21/IDE3	*hsd5 gal* (λ*c*1 *ts857 ind1 sam7 nin5 lacUV5*-T7 gene 1)	48
Plasmids		
pUC19	Ap^r, *lac* promoter/operator	53
M13.mp19	*lacZ′*	26
pRKTV14	Tn*903*::Tn*7*::Tn*5*	58
pUN121	λ*c*I, Ap^r, Tc^r [λ*c*I promoter]	31
pKT230	Sm^r, Km^r; RSF1010 derivative	4
pCR1000	*lacZ′* [T7 promoter], Km^r	Invitrogen
pPX7	*catB*::Tn*5*; pUC19 (Ap^r, Km^r) (*Sal*I partial from PRS3003)	This study
pPX9	*catC*::Tn*5*; pUC19 (Ap^r, Km^r) (*Sal*I partial from PRS3002)	This study
pPX30	*catB*::Tn*5′*; pUC19 (Ap^r, $CatC^+$) (*Hin*dIII subclone from pPX7)	This study
pPX31	*catC*::Tn*5′*; pUC19 (Ap^r, $CatB^+$) (*Pst*I subclone from pPX9)	This study
pPX63	*catR*::Tn*5*; pUC19 (Ap^r, Km^r) (*Sal*I partial from PRS3026)	This study
pPX66	*catR*; pUN121 (Ap^r, Km^r) (*Hin*dIII clone from PRS2000)	This study
pPX67	*catR*; pBR322 (Ap^r, Km^r) (*Sal*I subclone from pPX66)	This study
pPX72	*catR*; pKT230 (Ap^r, Km^r) (*Hin*dIII subclone from pPX66)	This study
pPX76	*catR*; pKT230 (Ap^r, Km^r) (*Sal*I-*Xho*I subclone from pPX67)	This study
pGHR1	*catR*; [T7 promoter]; pCR1000 (Km^r) (PCR clone from PRS2000)	This study

[a] Ap, ampicillin; Km, kanamycin; Tc, tetracycline; Ben, benzoate; Pob, *p*-hydroxybenzoate.

putida strains were grown to late exponential phase in 200-ml cultures; *E. coli* strains were grown overnight in 50 ml of L broth, diluted fourfold with L broth, and incubated for 4 to 8 h with isopropyl-β-D-thiogalactopyranoside (IPTG) added to a final concentration of 0.1 mM. DNA preparations were obtained from cells grown in L broth. Mobilization of pKT230 and its derivatives into *P. putida* was undertaken as previously described (44). For overexpression of CatR, 500-ml cultures of *E. coli* BL21/IDE3(pGHR1) were grown overnight in L broth (fortified with 50 μg of kanamycin per ml) and diluted 30-fold with L broth containing 0.1 mM IPTG.

Isolation of Tn*5* insertion mutants. Following overnight growth in L broth, 0.1 ml of *E. coli* Neco100 (harboring the suicide plasmid pRKTV14 bearing Tn*5*) (58) was mixed with 0.1 ml of *P. putida* PRS2000 on L-broth agar and grown overnight at 30°C. The resulting cells were suspended in 5 ml of basal medium and plated onto selective medium containing kanamycin, 0.5 mM succinate, and 5 mM benzoate so as to give approximately 100 colonies per plate. Following 3 days of incubation, small kanamycin-resistant (Km^r) colonies were examined for the ability to grow with *p*-hydroxybenzoate but not with benzoate. Mutants with this phenotype were further analyzed for evidence indicating insertion of Tn*5* within *cat* genes.

Purification and in vitro manipulation of DNA. Chromosomal DNA was isolated by the procedure of Berns and Thomas (5). Plasmid DNA was prepared by the clear lysate procedure (3) and, when necessary, was further purified over a cesium chloride-ethidium bromide gradient. Restriction endonuclease digestion and ligation of DNA were performed according to the manufacturer's specifications. Cloning of digested chromosomal DNA into *E. coli* expression vector pUC18 or pUC19 (53) and M13 phage DNA (26) was accomplished by using 10 μg of chromosomal DNA and 1 μg of plasmid or phage DNA. Vector DNA was digested and dephosphorylated according to published procedures (41). Transformation of *E. coli* was achieved by using either calcium chloride-treated cells (41) or cells prepared by the method of Hanahan (15). Inactivation of the *lacZ*-complementing gene by insertion of recombinant DNA in pUC or PCR1000 plasmids was detected by the failure of recombinant colonies to yield blue color on plates containing 0.1 mM IPTG and 2% (wt/vol) 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) in L-broth agar.

Measurement of enzyme activity. Cells were harvested by centrifugation at 10,000 × *g* for 5 min and resuspended in a 0.005 volume of 50 mM Tris-HCl (pH 7.5)–10% (vol/vol) glycerol–5 mM $(NH_4)_2SO_4$–5 mM $MgCl_2$–1 mM EDTA–1 mM dithiothreitol. Cell extracts were prepared by 15-s bursts of sonication with a Branson Electronics Sonifier, and cell debris was removed by centrifugation at 17,500 × *g* for 30 min. All manipulations were carried out at 4°C. The specific activities of enzymes were determined by established procedures (24, 30) and are expressed as percentages of those found in cells in which the enzymes had been fully induced by growth with 10 mM glucose in the presence of 30 mM adipate, a gratuitous inducer of *cat* gene expression in *P. putida* (35).

Chemicals and reagents. All restriction enzymes used in this study were obtained from commercial sources, as were the chemicals and enzymes used in the manipulation of DNA. *cis*,*cis*-Muconate and muconolactone were gifts from the Celgene Corporation. All other chemicals were of reagent grade or of the highest-quality grade available.

Primer extensions. A modification of previously described procedures (55) was used to characterize *cat* transcriptional initiation sites in total RNA isolated from *P. putida* PRS2000 grown under conditions that either did or did not induce expression of the *cat* genes. The purified RNA was pretreated for 11 min at 37°C in a 200-μl reaction mixture containing 9 mM Na_2HPO_4, 1 mM NaH_2PO_4, 200 mM NaCl, and 1 M formaldehyde. The reaction mixture was precipitated with 2 volumes of ethanol, centrifuged at 10,000 × *g* for 20 min, washed with 70% ethanol, dried under vacuum, and redissolved in distilled deionized water. Synthetic oligonucleotides, with sequences chosen to yield quantifiable reverse transcription products from the major transcripts within the *catR,BC,A* gene cluster, were radiolabeled by using [γ-^{32}P]ATP (Amersham) and T4 polynucleotide kinase (Boehringer) according to the manufacturers' specifications. The newly labeled primers were partially purified from unincorporated radiolabeled nucleotides by passage through G-25 micro-select spin columns (5′→3′ Inc.) and then used to prime the reverse transcription of the cellular RNA according to previously described methods (22). The synthetic oligonucleotide primers were 5′TCTGCCGGCTCAGGGCGGTTGGGCAATGTGC3′ (for *catR*) and 5′GCCTCACCGATACCTTCCACGCCATCGCTGC3′ (for *catB*).

Overexpression and partial purification of the *catR* regulatory gene. PCR technology was used to clone *catR* behind a T7 promoter (48) within PCR1000 (Invitrogen) (Table 1). The designed primers encompass nucleotide sequences flanking *catR* at its respective 5′ and 3′ ends. Noninvasive base substitutions (underlined in the following primer nucleotide sequences) were introduced into the synthetic DNA in order to minimize possible effects that CatR might exert on its own expression by binding to DNA upstream of its gene. Furthermore, the base substitutions (shown in boldface within the primer DNA sequences) incorporated restriction sites that facilitated the introduction of the PCR-amplified *catR* gene into plasmid pCR1000 in order to form pGHR1 (Table 1). The chemically synthesized primers were 5′TTT**AAGCTT**GGAGGTCTGATGGA

TABLE 2. Effects of Tn*5* insertions on expression of *cat* genes of *P. putida*[a]

Strain	Phenotype (genotype)	Growth conditions	Relative activity		
			MLE (*catB*)	MI (*catC*)	CO (*catA*)
PRS2000	Wild type	Adipate + glucose	100	100	100
		Glucose	<1	<1	4
PRS3002	Ben$^-$ Pob$^+$ (*catC102*::Tn*5*)	Adipate + glucose	93	<1	100
PRS3003	Ben$^-$ Pob$^+$ (*catB103*::Tn*5*)	Adipate + glucose	<1	<1	88
PRS3026	Ben$^-$ Pob$^+$ (*catR126*::Tn*5*)	Adipate + glucose	8	13	5
		Glucose	<1	<1	4

[a] Cells were grown with 10 mM glucose in the presence of 30 mM adipate (a nonmetabolizable inducer that elicits full expression of the *cat* genes) (35). Fully induced specific activities, similar to those found in benzoate-grown wild-type cells, were as follows: CatA, 1.0 U/mg of protein; CatB, 0.5 U/mg of protein; and CatC, 2.2 U/mg of protein. Uninduced levels of CatB and CatC were below the limit of detection; uninduced levels of CatA were 4% of the specific activities found in fully induced cells. Ben, benzoate; Pob, *p*-hydroxybenzoate; MLE, muconate lactonizing enzyme; MI, muconolactone isomerase; CO, catechol 1,2-dioxygenase.

GCTG3′ and 5′GG**GAATTC**GGTTTTACTCACACGATCTG3′. Positions corresponding to the primers are indicated in the nucleotide sequence of the *catR,BC,A* sequence shown in Fig. 3.

After expression of *catR* was elicited by growth of *E. coli* BL21/1DE3(pGHR1) in the presence of IPTG, cells were concentrated through a Millipore cell concentrator at 25 lb/in^2 and subsequently centrifuged for 10 min at 7,000 × *g*. The pellets were resuspended to a volume of 20 ml with storage buffer (50 mM Tris-HCl [pH 7.5], 0.5 mM EDTA, 1 mM dithiothreitol, 5% [vol/vol] glycerol). The cells were broken at 20,000 lb/in^2 in a French pressure cell (Amicon), and the resulting extract was centrifuged at 40,000 × *g* for 30 min. The supernatant was clarified by further centrifugation at 100,000 × *g* for 60 min and stored in 1-ml samples at −70°C.

DNA mobility shift assays. Mobility shift assays were used to assess the specific binding of CatR to DNA fragments containing either the *catR-catB* promoter region or *catC*. Using a previously described protocol (22), we used PCR to amplify both fragments from cloned chromosomal DNA from *P. putida* PRS2000. A 286-bp fragment encompassing the *catR-B* intergenic region was amplified with the following primers: 5′GGG**GAATTC**CGGGTGAAGTTCAG GGTCTC3′ and 5′TTT**GGATCC**TGCATTGCCAGCTTGTGCG3′ (see Fig. 3). Marked in boldface in these sequences are *Eco*RI and *Bam*HI sites included in the primer sequences to facilitate cloning of the fragment for further manipulation. The primers used to amplify a 411-bp fragment encompassing the *catC* gene were 5′GGCCTGACCCTGGACGAACAG3′ and 5′CACGGTCATCGG GCGTACCTC3′ (see Fig. 3). Fragments of the predicted sizes were yielded by each PCR amplification. Each fragment was radiolabeled according to the manufacturer's specifications, using [α-^{32}P]dATP and a random priming technique (9). Unincorporated nucleotides were removed by passage of the reaction mix through G-50 micro-select spin columns (Boehringer). The purified PCR fragments were used for mobility shift assays according to a modification of a published procedure (12). The reaction mixes contained, in 20 μl, 1 μg of poly(dI-dC) (Pharmacia), 10,000 cpm of probe DNA, and 2 μl of 10× binding buffer (containing 500 mM Tris-HCl [pH 7.5], 50% [vol/vol] glycerol, 0.5 mM EDTA, and 10 mM dithiothreitol). The mixture of protein and DNA was allowed to equilibrate for 30 min at 30°C prior to being loaded onto a prerun 5% TBE (45 mM Tris-borate, 1 mM EDTA [pH 7.6]) polyacrylamide gel that was then run at 25 mA. After an appropriate time interval, the gel was dried and an autoradiograph of the gel was prepared.

Nucleotide sequence analysis. DNA sequences were determined by a modification of the dideoxy-chain termination method (42), using enzymes and reagents supplied by United States Biochemical Corporation (Cleveland, Ohio), excepting ^{35}S-dATP, which was purchased from Amersham, Inc. (Arlington, Ill.). The sequence of the entire *catR,BC,A* region (along with the localization of Tn*5* insertion within inactivated genes) was determined on both strands, using a range of double-stranded plasmid and single-stranded M13 phage subclones. Owing to the relatively high G+C content of *Pseudomonas* DNA, each region of DNA was sequenced in stretches of less than 300 bases, and these sequences were confirmed by additional sequence analysis using deaza-dGTP from United States Biochemical. When convenient sites for restriction were not readily available, oligonucleotide primers, typically 18 to 23 bp in length, were synthesized. These primers were used with either plasmid or phage templates. Interpretation of nucleotide sequence was assisted by the MacVector computer analysis program, version 4.1, from International Biotechnologies (New Haven, Conn.).

Nucleotide sequence accession number. The sequence containing the *catR, BC,A* region from *P. putida* PRS2000 is available in GenBank under accession number U12557.

RESULTS

Isolation of Tn*5* mutants of PRS2000. Mutagenesis of strain PRS2000 with Tn*5* yielded mutant strains that exhibited resistance to kanamycin, inability to grow with benzoate, and ability to grow at the expense of *p*-hydroxybenzoate. Three isolates, PRS3002, PRS3003, and PRS3026, were chosen for further study (Table 1). Southern blotting of restriction digests (with the internal 2.7-kbp *Bgl*II fragment of Tn*5* as a probe) confirmed that each of these mutant strains had acquired a single copy of Tn*5*.

Exposure of wild-type *P. putida* cells to adipate resulted in full expression of *catB*, *catC*, and *catA* (Table 2). The mutation in strain PRS3002 prevented expression of *catC*; the mutation in strain PRS3003 prevented expression of *catB* and *catC* (Table 2). Strain PRS3026 exhibited the phenotype of a *catR* mutant (54) by failing to express any of the *cat* structural genes at induced levels (Table 2). Subsequent nucleotide sequence analysis revealed the location of Tn*5* within *catR* in PRS3026, within *catB* in PRS3003, and within *catC* in PRS3002 (Fig. 2 and 3). As the gene order *catBC,A* was elucidated, it became apparent that *catB103*::Tn*5* exerts a polar effect preventing transcription of the directly downstream *catC*. Expression of *catA*, which lies directly downstream from *catC*, is unimpeded by the polar mutation (Table 2). This evidence suggests that an additional promoter contributes to transcription of *catA*.

Cloning of *catB*. A partial *Sal*I digest of chromosomal PRS3002 DNA was ligated into the *Sal*I site of pUC19, and the ligated material was introduced into *E. coli* JM109 and subjected to selection for Km^r. A selected plasmid, designated pPX9, carried an insert of about 10 kbp and expressed *catB* but not *catA* or *catC* when the host *E. coli* cells were exposed to IPTG. Nucleotide sequence analysis of the subclone pPX31 revealed that pPX9 contained *catR*, *catB*, a portion of *catC*, and none of *catA* (Fig. 2 and 3).

Cloning of *catC* and *catA*. Similar procedures were used to recover a recombinant pUC19 plasmid (pPX7) containing a 25-kbp partial *Sal*I restriction fragment from PRS3003. Exposure of *E. coli* JM109(pPX7) to IPTG resulted in expression of *catC* but not *catB*. Nucleotide sequence analysis of the subclone pPX30 revealed that pPX7 contains all of *catC*, a portion of *catB*, and the complete *catA* structural gene (Fig. 2 and 3).

Cloning of *catR*. The phenotype of PRS3026 (Table 2) suggested that it was blocked in *catR*, a transcriptional activator gene known to lie near *catBC* (54). To prepare a probe for the *catR,BC* chromosomal region, a *Sal*I digest of DNA from strain PRS3026 was used to obtain DNA containing both Km^r from the inserted Tn*5* transposon and a portion of chromosomal DNA flanking the site of the transpositional insertion. A *Sal*I digest of DNA from strain PRS3026 was introduced into pUC19 by ligation, and this material was transformed into *E. coli* JM109 and subjected to selection for both ampicillin resistance and Km^r. This procedure yielded a recombinant plasmid (pPX63) containing 2.7 kbp of Tn*5* DNA and 0.8 kbp of DNA from the *P. putida* chromosome (Fig. 2).

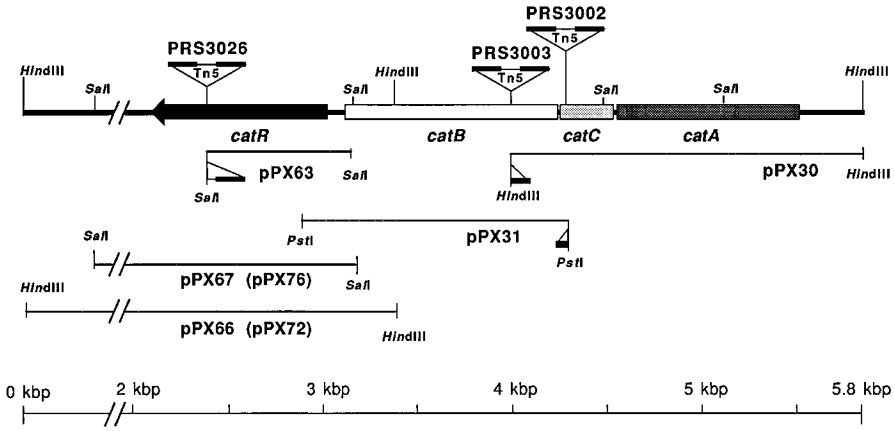

FIG. 2. Schematic representation of the *catR,BC,A* cluster from the *P. putida* biotype strain PRS2000. Depicted are sites of Tn*5* insertion in chromosomal DNA from mutant strains and cloned DNA segments within recombinant plasmids isolated in the course of the investigation.

DNA from pPX63 was used as a probe to detect DNA from the *catR* region of the wild-type *P. putida* chromosome. A *Hin*dIII restriction digest of this DNA was prepared and used to form ligation recombinants containing restriction fragments with the *Hin*dIII site of λ*c*I within a pUN121 vector (31). This plasmid contains a tetracycline resistance (Tc^r) gene which is repressed by the λ*c*I gene product, and thus insertion of DNA into λ*c*I results in expression of the Tc^r gene, giving a positive selection for cloned DNA. Screening of 5,000 *E. coli* JM109 colonies that had expressed Tc^r after transformation yielded one strain that hybridized with DNA from pPX63. Restriction mapping of the plasmid (pPX66) from this transformant revealed a 3.4-kbp *Hin*dIII insert containing *catR*. A subcloned 3.0-kbp *Sal*I fragment was inserted into pBR322 to form pPX67 (Fig. 2), and this plasmid was used for further restriction mapping, hybridization analysis, subcloning, and DNA sequencing.

Complementation of a known mutation in *P. putida catR*. Earlier genetic analysis had indicated the location of the *catR1122* mutation within *P. putida* PRS2014 (54). Presence of the wild-type *catR* gene in pPX66 and pPX67 was confirmed by complementation of *catR1122* by restriction fragments carried in these plasmids. The *Hin*dIII fragment from pPX66 and the *Sal*I fragment from pPX67 were introduced into the respective *Hin*dIII and *Xho*I sites of pKT230, yielding the recombinant plasmids pPX72 and pPX76 (Table 1; Fig. 2), which were selected after transformation into *E. coli* JM109. Mobilization of the plasmids into strain PRS2014 (*catR1122*) yielded recombinant strains that grew with benzoate. Repeated efforts to obtain the recombinant plasmids from these strains yielded only plasmids that were physically indistinguishable from pKT230. Thus, it appears that recombination removed the wild-type DNA carried in the plasmids as it replaced the chromosomal region containing the mutation. This finding accounts for difficulties repeatedly encountered in efforts to recover wild-type *P. putida* genes on recombinant plasmids that had been selected on the basis of their ability to complement mutations in the *P. putida* biotype strain.

DNA sequence analysis of the *catR,BC,A* cluster. (i) Conserved nucleotide sequences. The nucleotide sequence of the *catR,BC,A* cluster (Fig. 3) was constructed by analysis of overlapping fragments of cloned DNA (Fig. 2). Presence of four open reading frames was confirmed by computer analysis weighted to accommodate frequently used codons within the sequence of *Pseudomonas* genes. Predicted sizes of the translated polypeptide sequences are in agreement with reported molecular weights for CatR from *Pseudomonas* strain RB1 (40), CatB and CatC from *P. putida* PRS1 (14, 17), and CatA from *A. calcoaceticus* (29). Identification of the sites of Tn*5* insertion (Fig. 2 and 3) confirmed the genotypes that had been inferred on the basis of enzymatic analysis of mutant strains (Table 2).

Three genes (*catR*, *catB*, and *catC*) from the *P. putida* biotype strain closely resemble their counterparts from *Pseudomonas* strain RB1. The identity of the aligned nucleotide sequences for these genes approximates 90% (Fig. 3). Furthermore, aligned regions between *catC* and *catA* reveal only five substitutions over 42 nucleotides (Fig. 3). Adjustment by removal of a single nucleotide from the published *Pseudomonas* strain RB1 sequence allows its alignment with the *P. putida catA* nucleotide sequence, with identity of 90% for the 180 residues known in both sequences (Fig. 3). The predicted amino acid sequence of *P. putida* CatA shares 52% amino acid sequence identity with catechol 1,2-dioxygenase from *A. calcoaceticus* (29) and also resembles the primary structure of other enzymes that share this function (8, 19). Thus, it is apparent that *catA*, which is separated from *catC* by several kilobases in the closely related species *A. calcoaceticus* and *P. aeruginosa*, has been selected in a position directly downstream from *catC* in *Pseudomonas* strain RB1 and *P. putida* PRS2000.

(ii) Divergent nucleotide sequences. Extensive divergence has occurred at two locations in the *catR,BC,A* regions of the two *Pseudomonas* strains. One location is directly downstream from *catR*, where similarity of nucleotide sequence ends abruptly (Fig. 3). The other location is the *catB-catC* intergenic region, which extends 21 nucleotides in *P. putida* and 138 nucleotides in *Pseudomonas* strain RB1. As shown in Fig. 4 and 5, these two locations are marked by the presence of a differently positioned 85-bp element which is in some respects analogous to the repetitive extragenic palindromic (REP) sequences of enteric bacteria (7, 13, 47).

In other portions of *catR,BC,A*, base substitutions distinguishing the *Pseudomonas* strain RB1 and *P. putida* nucleotide sequences tend to be fairly evenly distributed. An exception to this pattern occurs directly upstream from *catB*, where a cluster of nine base pair differences (substitutions or single base pair insertions) in the *Pseudomonas* strain RB1 sequence disrupt a 30-nucleotide stem-loop structure that appears to have been selected in *P. putida* (Fig. 4). As described below, primer extension analysis indicates that the stem-loop structure occu-

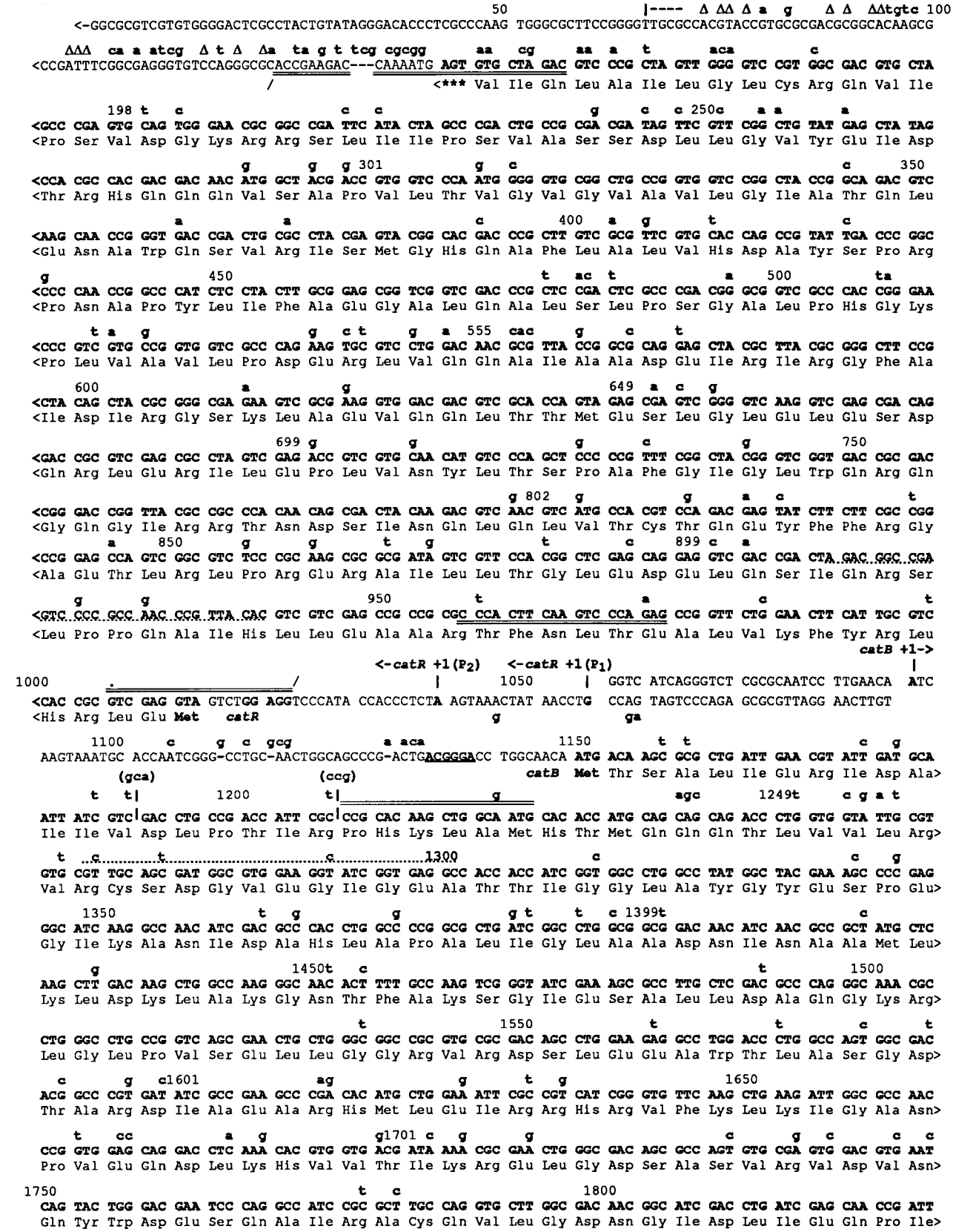

FIG. 3. Nucleotide sequence of the coding strands for the divergently expressed *catR,BC,A* region from the *P. putida* biotype strain PRS2000. Deduced amino acid sequences are shown below the corresponding DNA sequences, which are presented in boldface. Also shown in boldface are nucleotides identified by primer extension as transcriptional start sites, which are further defined by +1 and reference to the appropriate gene. The DNA sequences of oligonucleotide primers used for amplification of specific DNA segments are indicated by either single or double lines, with the direction of the primer indicated by its position relative to the *catBC,A* coding sequence; underlining marks primers extending in the direction of *catB* transcription, whereas lines above the sequence refer to primers extending in thedirection opposite that of *catB* transcription. Dotted lines above or below the DNA sequence indicate the structure of synthetic oligonucleotides used for identification of transcriptional starts by primer extension with RNA from induced and uninduced cells. The annotation Tn*5* designates sites where insertion of the

```
                    a c1850       a                    gc   g                    c   1900g
TCG CGT ATC AAC CGC GGT GGC CAG GTG CGC CTG AAC CAG CGC ACG CCT GCA CCG ATC ATG GCC GAT GAG TCC ATC GAA AGC
Ser Arg Ile Asn Arg Gly Gly Gln Val Arg Leu Asn Gln Arg Thr Pro Ala Pro Ile Met Ala Asp Glu Ser Ile Glu Ser>

          c                                    1950        c          a                    g
GTC GAG GAT GCC TTC AGC CTG GCT GCC GAC GGC GCG GCC AGC ATC TTT GCC CTG AAA ATC GCC AAG AAC GGC GGC CCG CGT
Val Glu Asp Ala Phe Ser Leu Ala Ala Asp Gly Ala Ala Ser Ile Phe Ala Leu Lys Ile Ala Lys Asn Gly Gly Pro Arg>

  c   a2000   t   c              t              c     c  t        g   2050              a
GCT GTG CTG CGC ACT GCG CAA ATC GCC GAG GCC GCC GGT ATC GGC CTG TAT GGC GGC ACC ATG CTC GAG GGC TCG ATC GGC
Ala Val Leu Arg Thr Ala Gln Ile Ala Glu Ala Ala Gly Ile Gly Leu Tyr Gly Gly Thr Met Leu Glu Gly Ser Ile Gly>

          t       t   t           g2102 c        g                              g           2150
ACC CTG GCC TCG GCC CAC GCC TTC CTT ACC TTG CGC CAA CTG ACC TGG GGC ACC GAG CTG TTC GGC CCG CTG CTG CTG ACC
Thr Leu Ala Ser Ala His Ala Phe Leu Thr Leu Arg Gln Leu Thr Trp Gly Thr Glu Leu Phe Gly Pro Leu Leu Leu Thr>

                                  c                    2200        a
GAA GAA ATC GTC AAC GAG CCG CCG CAA TAT CGC GAC TTC CAG CTG CAC ATT CCC CGT ACC CCA GGC CTG GGC CTG ACC CTG
Glu Glu Ile Val Asn Glu Pro Pro Gln Tyr Arg Asp Phe Gln Leu His Ile Pro Arg Thr Pro Gly Leu Gly Leu Thr Leu>

                  2250                          <---138 bp insert--->        2300
GAC GAA CAG CGC CTG GCG CGC TTC GCC CGT CGC TGA <TCAGCAAACAAGGAGATAATC> ATG CTG TTC CAC GTG AAG ATG ACC GTG
Asp Glu Gln Arg Leu Ala Arg Phe Ala Arg Arg ***>                catC  Met Leu Phe His Val Lys Met Thr Val>

              c                          2350c g   g              c                          g   a
AAA CTG CCG GTG GAC ATG GAC CCG GCC AAG GCT ACC CAG CTC AAG GCC GAT GAA AAG GAA CTG GCC CAG CGC CTG CAA CGT
Lys Leu Pro Val Asp Met Asp Pro Ala Lys Ala Thr Gln Leu Lys Ala Asp Glu Lys Glu Leu Ala Gln Arg Leu Gln Arg>

2400    c   t                      g      agtC c          c    2450                      a t
GAA GGG ACT TGG CGC CAC CTG TGG CGC ATT GC| GGG CAC TAC GCA AAC TAC AGC GTG TTC GAT GTG CCC AGC GTC GAG GCG
Glu Gly Thr Trp Arg His Leu Trp Arg Ile Ala Gly His Tyr Ala Asn Tyr Ser Val Phe Asp Val Pro Ser Val Glu Ala>

 gc a                 2500g         c                    t          a       t       t  2550 c
TTG CAC GAC ACC CTG ATG CAA CTG CCG TTG TTC CCG TAC ATG GAC ATC GAG GTC GAC GGC CTG TGC CGG CAT CCC TCG TCG
Leu His Asp Thr Leu Met Gln Leu Pro Leu Phe Pro Tyr Met Asp Ile Glu Val Asp Gly Leu Cys Arg His Pro Ser Ser>

  t                       gt            2600                    c a   a           t
ATC CAC AGC GAC GAC CGC TGA TACGCACC TGTCTACCTG ACAAGAACAA TATGAGGTAC GCCCG ATG ACC GTG AAA ATT TCC CAC ACT
Ile His Ser Asp Asp Arg ***>                    (g)                catA  Met Thr Val Lys Ile Ser His Thr>

2650    t g    g   g          g g        c        | c    t   t2700
GCC GAC ATT CAA GCC TTC TTC AAT AAA GTG GCG GGC CTG|GAC CAC GCC GAA GGC AAC CCA CGC TTC AAG CAG ATC ATC CTG
Ala Asp Ile Gln Ala Phe Phe Asn Lys Val Ala Gly Leu Asp His Ala Glu Gly Asn Pro Arg Phe Lys Gln Ile Ile Leu>

                g      t   a  2755       g                                          t   2800c ---|
CGT GTG CTG CAA GAC ACC GCG CGC CTG GTC GAA GAC CTG GAA ATC ACC GAA GAC GAG TTC TGG CAC GCC ATT GAT TAT CTC
Arg Val Leu Gln Asp Thr Ala Arg Leu Val Glu Asp Leu Glu Ile Thr Glu Asp Glu Phe Trp His Ala Ile Asp Tyr Leu>

                                          2850
AAC CGC CTG GGT GGC CGT AAC GAA GCC GGC CTG CTG GCT GCA GGC CTG GGC ATC GAG CAC TTC CTC GAC CTG CTG CAG GAC
Asn Arg Leu Gly Gly Arg Asn Glu Ala Gly Leu Leu Ala Ala Gly Leu Gly Ile Glu His Phe Leu Asp Leu Leu Gln Asp>

      2900                                                          2950
GCC AAG GAT GCC GAA GCC GGC CTG GGT GGC GGC ACC CCG CGT ACC ATC GAA GGC CCG CTG TAT GTC GCC GGG GCG CCG CTG
Ala Lys Asp Ala Glu Ala Gly Leu Gly Gly Gly Thr Pro Arg Thr Ile Glu Gly Pro Leu Tyr Val Ala Gly Ala Pro Leu>

                                  3000                                                          3050
GTA CAA GGC GAA CGG CGC ATG GAC GAC GGC ACC GAC CCA GGG GTG GTG ATG TTC CTG AAA GGC CAG GTG TTC GAC GCC GAA
Val Gln Gly Glu Arg Arg Met Asp Asp Gly Thr Asp Pro Gly Val Val Met Phe Leu Lys Gly Gln Val Phe Asp Ala Glu>

                                                  3100
GGC AAG CCG CTG GCC GGC GCC ACC GTC GAC CTC TGG CAC GCC AAC ACC CAA GGT ACT TAT TCG TAC TTC GAC TCG ACC CAG
Gly Lys Pro Leu Ala Gly Ala Thr Val Asp Leu Trp His Ala Asn Thr Gln Gly Thr Tyr Ser Tyr Phe Asp Ser Thr Gln>

                3150                                                          3200
TCC GAG TAC AAC CTG CGC CGC CGC ATC ATC ACC GAT GCC GAG GGC CGC TAC CGT GCG CGC TCC ATC GTG CCA TCG GGG TAT
Ser Glu Tyr Asn Leu Arg Arg Arg Ile Ile Thr Asp Ala Glu Gly Arg Tyr Arg Ala Arg Ser Ile Val Pro Ser Gly Tyr>

                                          3250
GGC TGC GAC CCG CAG GGC CCG ACC CAG GAA TGC CTG GAC TTG CTC GGG CGT CAT GGC CAG CGC CCG GCG CAT GTG CAC TTC
Gly Cys Asp Pro Gln Gly Pro Thr Gln Glu Cys Leu Asp Leu Leu Gly Arg His Gly Gln Arg Pro Ala His Val His Phe>

3300                                                          3350
TTC ATC TCG GCA CCC GGG CAC CGC CAT CTG ACC ACG CAG ATC AAC TTC GAG GGT GAC AAG TAC CTG TGG GAC GAC TTT GCC
Phe Ile Ser Ala Pro Gly His Arg His Leu Thr Thr Gln Ile Asn Phe Glu Gly Asp Lys Tyr Leu Trp Asp Asp Phe Ala>

                          3400                                                          3450
TAC GCC ACC CGT GAC GGG CTG ATC GGC GAG CTG CGT TTT GTC GAG GAT GCT GCG GCG GCG CGT GAC CGC GGC GTT CAA GGC
Tyr Ala Thr Arg Asp Gly Leu Ile Gly Glu Leu Arg Phe Val Glu Asp Ala Ala Ala Ala Arg Asp Arg Gly Val Gln Gly>

                                          3500
GCG CGT TTT GCC GAA CTG GCG TTC GAC TTC CAC CTG CAG GGG GCC ACG GCG GTG GAG GCC GAG GCA CGC AGC CAC CGG CCG
Ala Arg Phe Ala Glu Leu Ala Phe Asp Phe His Leu Gln Gly Ala Thr Ala Val Glu Ala Glu Ala Arg Ser His Arg Pro>

        3550                                                          3600
CGT GCG TTG CAG GAA GGC TGA AACCGTTGA GGCTGCTGTG CAGCCGTTCG CGGCACAAGG CCGCGAAAGG GGCGCACAGC GCCCAAAATG
Arg Ala Leu Gln Glu Gly ***>

                  3650
CTGACCATTT TTCGCTGTGT TTTTGCCTGG GCCTTTCCGT GCGGGGAGGG CCCGGG->
```

transposon have been documented in mutant genes. Shown in lowercase above the DNA sequence are nucleotide substitutions and insertions present in the sequence reported for *Pseudomonas* strain RB1, with | – – – – and – – – – | used to locate the extremities of the DNA that has been sequenced from this strain (1, 2, 39). The symbol △ indicates nucleotides that are absent in the aligned *Pseudomonas* strain RB1 sequence. This sequence contains a 138-bp insert at the indicated position between *catB* and *catC*.

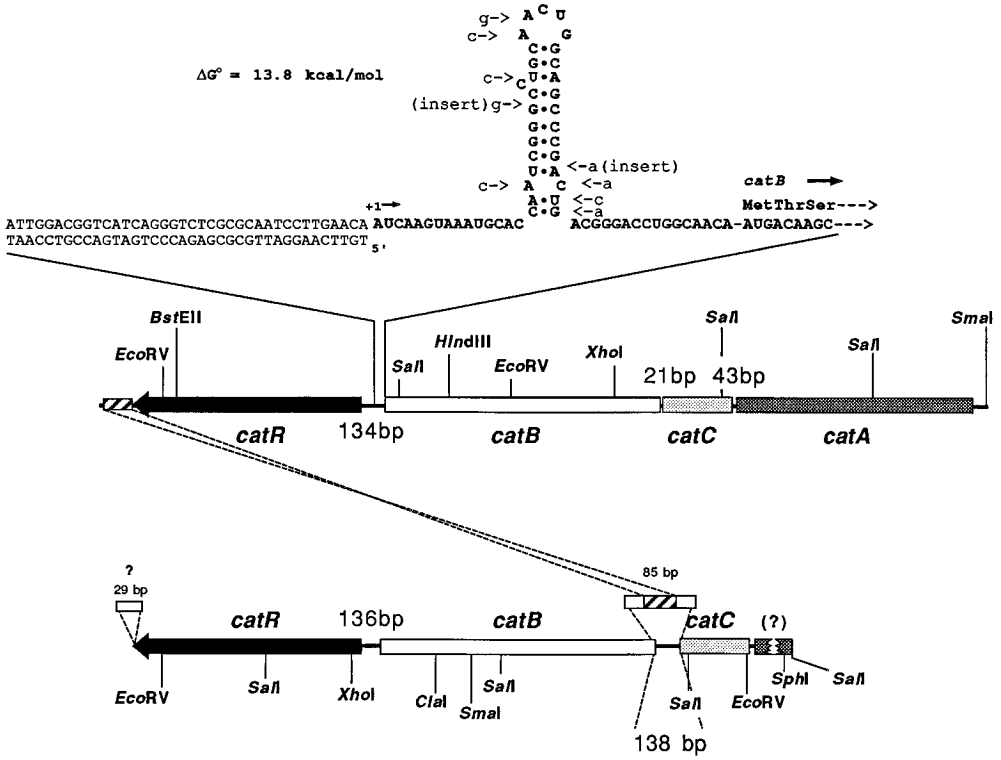

FIG. 4. Comparison of the overall organization of *cat* genes in the *P. putida* biotype strain PRS2000 and in *Pseudomonas* strain RB1. The *catB* and *catC* genes are separated by 21 bp of DNA in the former strain, whereas 138 bp of DNA lie between the genes in strain RB1. Within this 138-bp region is an 85-bp segment (striped box) that closely resembles DNA lying directly downstream from the end of *catR* in *P. putida* biotype strain PRS2000. It is in these two regions that the *cat* DNAs from the two *Pseudomonas* strains exhibit extreme sequence divergence (Fig. 3). Also shown is a potential RNA stem-loop structure near the beginning of the *catB* transcript in the *P. putida* biotype strain. Nucleotide base substitutions which are apparently localized within this region in the corresponding transcript from *Pseudomonas* strain RB1 (1) tend to severely disrupt this structure and are designated in lowercase as bases whose positions in the PRS2000 sequence are indicated by arrows. One kilocalorie = 4.184 kJ.

pies a position that would allow it to modulate *catR,BC* transcription in *P. putida*.

Primer extension analysis. To identify the 5′ termini of the *catR* and *catBC* transcripts, a set of oligonucleotide primers, greater than 30 bp in length, was designed to hybridize specifically with *catR-B* RNA purified from *P. putida* PRS2000 cells grown in either the presence or the absence of the inducing growth substrate, benzoate. Reverse transcriptional analysis from these primers revealed that the *catR-B* promoter/operator region contains an abnormally high degree of secondary structure which resulted in the presence of numerous bands corresponding to points of termination of transcription. This background was not removed by variation in hybridization conditions but was overcome by preincubation of total RNA with 1 M formaldehyde. This treatment, first reported by Lodish (21), was found to be useful in reducing RNA secondary structure in order to determine its role in regulating translation of bacteriophage f2. In the present case, the pretreatment reduced background and significantly enhanced the relative intensity of the band corresponding to the 5′ end of the *catB* transcript elicited by growth of cells under inducing conditions (Fig. 6A). Use of a primer specific for *catB* revealed a major band, corresponding to the A at position 1088 of the nucleotide sequence (Fig. 3), that was formed only by RNA from induced cells (Fig. 6A). This location for the transcriptional start of *catB* corresponds to the site determined with *Pseudomonas* strain RB1. A second band, corresponding to the A at position 1097, also was observed (Fig. 6A). This band may represent the location of the 5′ end of an additional transcript or may be due to some partial degradation of *catB* RNA stalled at the base of the potential stem-loop structure near the 3′ side of the transcriptional start site (Fig. 5). This band has no counterpart in *Pseudomonas* strain RB1 (1), which lacks the nucleotide sequence that would allow for the stem-loop structure in the mRNA.

To determine the 5′ end of the *catR* transcript, a primer which hybridized specifically to this gene was used to reverse transcribe total cellular RNA from *P. putida* PRS2000. Bands corresponding to these transcriptional starts exhibited substantially lower intensity than those produced for *catB* transcripts in induced cells (Fig. 6). A *catR* transcriptional start site was clearly evident (Fig. 6B) at a location corresponding to A at position 1040 (Fig. 3); this position matches the transcriptional start reported for *catR* in *Pseudomonas* strain RB1 (40). Induced *P. putida* PRS2000 cultures contained an additional band (Fig. 6B) corresponding to the G base at position 1056 (Fig. 3). This band has no apparent counterpart in *Pseudomonas* strain RB1 and, in the absence of any discernible secondary structure within this region, may reflect the position of an additional transcriptional start in the *P. putida* biotype strain.

Dual binding capacity of the *catR* gene product. Use of a T7-6 expression vector (carrying a *lacZ′* multiple cloning site downstream from a phage T7 promoter) (49) permitted overexpression of the cloned *catR* regulatory gene which had been modified to independence of autoregulation. The resulting *catR* gene product was partially purified in a manner similar to

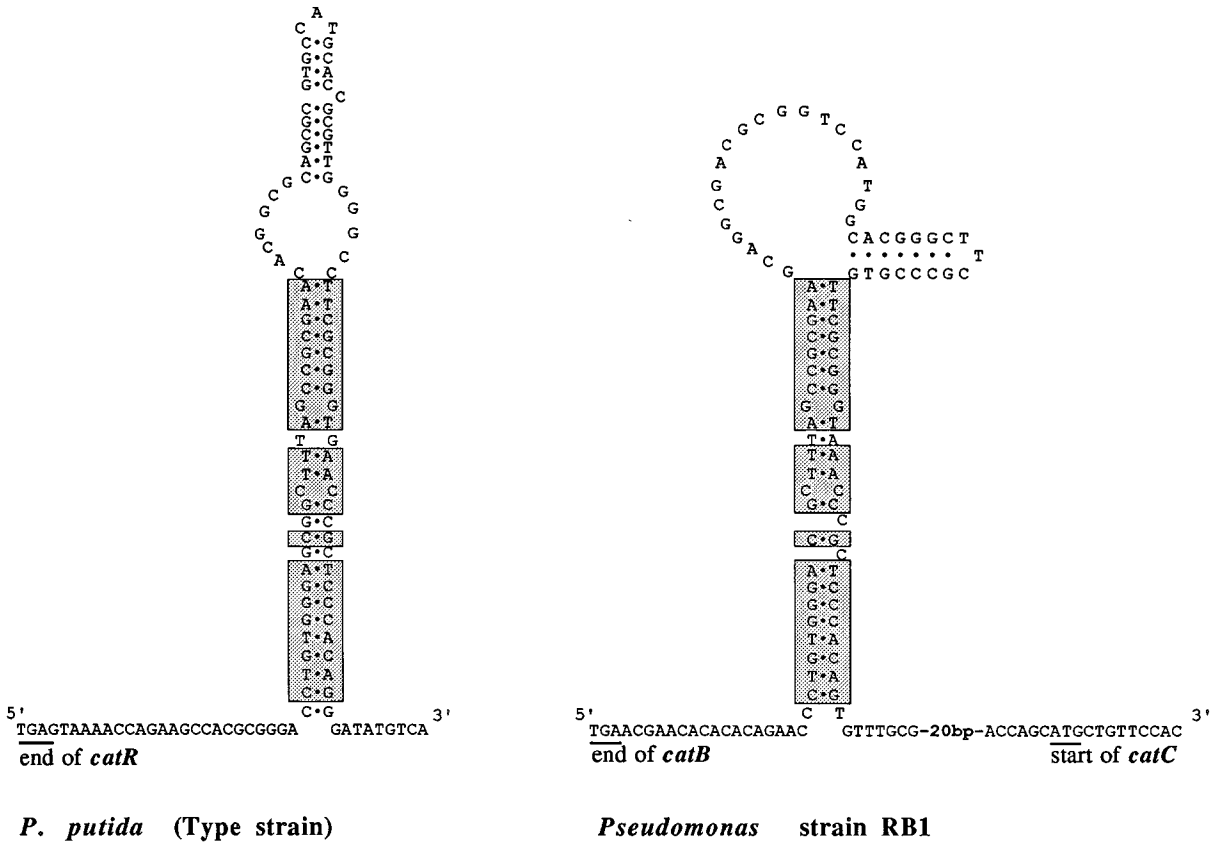

FIG. 5. Potential secondary structures in migratory DNA elements found downstream from *catR* in the *P. putida* biotype strain and downstream from *catB* in *Pseudomonas* strain RB1. Shaded residues have the potential ability to form a stem-loop structure that is conserved in both sequences. Additional stem-loop structures distinguish the two structures.

that described for its counterpart from *Pseudomonas* strain RB1 (40). Throughout all stages of purification, the presence of CatR was monitored by its ability to retard mobility of a 286-bp fragment that fully encompasses the *catR-B* promoter region (covering bases 957 to 1226; Fig. 3). Figure 7 demonstrates one such assay, in which the mobility of *catR-B* promoter fragment was retarded by partially purified CatR protein. Also revealed in Fig. 7 is the ability of a similar amount of CatR to bind specifically to a different DNA fragment containing the *catC* structural gene. Thus, CatR possesses the ability to bind to two physically separate regions in the *catR,BC,A* gene cluster. Interactions at one binding site appear to trigger expression of *catBC*, and it is likely that further analysis of the CatR binding site within the fragment containing *catC* may reveal specific interactions leading to independent expression of *catA*.

DISCUSSION

Loss of *P. putida* DNA inserts from plasmids. Crystalline CatB (14) and CatC (17) were purified from the *P. putida* biotype strain derivative PRS2000 (24), and optimal alignment of deduced amino acid sequence with three-dimensional structure demanded determination of the nucleotide sequences of the *catBC* region from this strain. A seemingly direct approach to cloning *catBC* was their selection in a broad-host-range plasmid by demanding complementation of a representative of the strain which carries a characterized deletion in *catB* (54). Such efforts met with frustration because clones that acquired *catB* from donor libraries contained only plasmids of the same size as the original vector. An explanation for this finding is that *P. putida* DNA is lost from the plasmid through recombination when carried in the presence of homologous chromo-

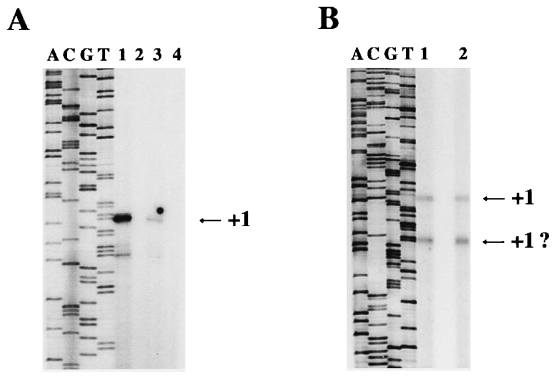

FIG. 6. Transcriptional start sites for *catB* and *catR*. The sites were determined by primer extension of RNA derived from cultures of *P. putida* biotype strain that have been induced (grown with benzoate and glucose) and uninduced (grown with glucose alone). (A) The start of the *catB* transcript. Lane 1, formaldehyde-treated RNA from induced cells; lane 2, formaldehyde-treated RNA from uninduced cells; lane 3, untreated RNA from induced cells; lane 4, untreated RNA from uninduced cells. (B) The start of the *catR* transcript determined with formaldehyde-treated RNA. Lane 1, RNA from induced cells; lane 2, RNA from uninduced cells.

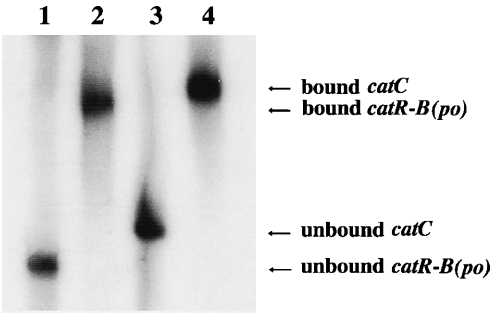

FIG. 7. Binding of CatR to separate DNA fragments from the *catR,BCA* region. Gel retardation demonstrating the specific affinity of 3-μg samples of partially purified CatR for either a 286-bp DNA fragment containing the *catR-catB* intergenic region (lanes 1 and 3) or a 390-bp DNA fragment containing the *catC* gene (lanes 3 and 4).

somal DNA. This interpretation is strengthened by observation of the apparent loss of the DNA inserts from pPX72 and pPX76 after selection for their function in PRS2014 (*catR1122*).

Evolutionary divergence of the *P. putida* biotype strain and *Pseudomonas* strain RB1. *Pseudomonas* strain RB1 resembles the *P. putida* biotype strain PRS2000 in the *catR,BC,A* gene arrangement, which is conserved in the two strains. In contrast, *catA* is separated by several kilobases from *catBC* in *A. calcoaceticus* and *P. aeruginosa*. Rearrangement of members of the *catA* and *catB* gene families is further demonstrated by plasmid-borne chlorocatechol genes *clcA* (homologous to *catA*) and *clcB* (homologous to *catB*), which are part of the *clcABD* cluster (10).

A difference between *P. putida* PRS2000 and *Pseudomonas* strain RB1 is indicated by the stability of DNA inserts carrying *cat* genes from strain RB1 when contained in plasmids replicating in strain PRS2000. This stability could be attributed most simply to nucleotide sequence differences sufficient to deter homologous recombination. Comparison of *catR*, *catB*, *catC*, and the available portion of *catA* indicates nucleotide sequence divergence of about 10%, with predominately silent base substitutions scattered fairly evenly through the genes (Fig. 3). This frequency of nucleotide divergence is similar to the frequency of base substitution in the aligned sequences of the respective *glpFKX* genes from *E. coli* and *Shigella flexneri* (50). Thus, nucleotide sequence divergence in the *catR,BC,A* in the two *Pseudomonas* strains is similar to that of the *glpFKX* region in cell lines that have been accorded different generic status in the enteric bacteria. It is noteworthy that recombination does not present an obstacle to maintenance of plasmid-borne *S. flexneri glpFKX* in the presence of *E. coli glpFKX* in the chromosome (38, 50).

Analogs of REP sequences in the *Pseudomonas cat* region. Dramatic divergence of the *glp* region in the two genera of enteric bacteria was achieved by two REP sequences (47) which lie between *glpK* and *glpX* in the chromosome of *S. flexneri* but not of *E. coli*. The REP sequences of the enteric bacteria are members of a conserved family of genetic elements that occur in transcribed, untranslated DNA segments. It is estimated that more than 500 REP sequences are contained in the *E. coli* genome (7).

Abrupt divergence of the *Pseudomonas catR,BC,A* nucleotide sequences in *P. putida* PRS2000 and *Pseudomonas* strain RB1 was achieved by insertion of similar palindromic genetic elements at different locations (Fig. 5). Like the REP sequences of the enteric bacteria, the *Pseudomonas* palindromic elements are inserted in transcribed, untranslated regions. The *Pseudomonas* palindromic sequences are 85 bp in length and thus are larger than the REP sequences, which range between 30 and 40 bp. On the other hand, the *Pseudomonas* palindromic units are smaller than the approximately 125-bp enteric repetitive intergenic consensus (ERIC) sequences (16, 43, 52) that are positioned differently in the *lux* regions of different luminescent bacteria (25). The *Pseudomonas* palindromic sequences bear no discernible sequence similarity to either REP of ERIC sequences, but this is not remarkable given the divergence of these sequences even within the enteric bacteria (7, 16). The *Pseudomonas* palindromic sequences may serve functions analogous to those served by REP and ERIC sequences. This possibility could be explored by examination of the distribution of the *Pseudomonas* palindromic sequences within the chromosome of the *P. putida* biotype strain and, more generally, among fluorescent *Pseudomonas* species.

Control of *catA* expression by *catR* in the *P. putida* biotype strain. Insertion of Tn*5* into *catR* blocks induced expression of both *catBC* and *catA* in *P. putida* PRS3026 (Table 2). This evidence strengthens the earlier conclusion, based upon transduction of spontaneous mutants, that *catR* encodes a transcriptional activator of the *cat* structural genes (54). In the present study, a fragment from the mutated gene was used as a probe to recover wild-type *catR*, and this gene was used to complement the *catR1122* mutation used in the original transductional analysis (54).

The properties of *P. putida* PRS3003 illustrate that insertion of Tn*5* into *catB* prevents transcription of the downstream *catC* gene (Table 2). This polar effect does not prevent inducible expression of the more distantly downstream *catA* in strain PRS3003 (Table 2). Thus, *catR* exerts pleiotropic effects and can apparently trigger *catA* expression independently of the control exerted upstream of *catB*. This inference is strengthened by a demonstration that CatR binds specifically to a DNA fragment containing *catC* as well as to the *catR-B* intergenic region (Fig. 7). If, as these observations suggest, CatR has the ability to activate *catA* expression in *P. putida* by binding directly upstream of this gene, a single duplication could account for much of the genetic change required for activation of *clcABD* by binding of ClcR (homologous with CatR) upstream of *clcA* (homologous with *catA*) (6).

Continuing possibility for interaction between CatR and the *clc* promoter is indicated by the ability of the *P. putida* CatR protein to complement a *clcR* null mutation (23). Avoidance of transcriptional activation of *catBC* by ClcR would require divergence of nucleotides that allow activation of *catBC* by ClcR in *P. putida*. In this regard, it is noteworthy that the *Pseudomonas* strain RB1 chromosome contains clustered mutations that disrupt a stem-loop RNA structure that has been selected at the beginning of the *catB* transcript in *P. putida*. The absence of this dyad symmetry in RB1, apparently reinforced by discrete sequence changes in this region, suggests that it could provide in *P. putida* a regulatory function that may be absent in *Pseudomonas* strain RB1.

Evolutionary divergence of the *cat* region in fluorescent pseudomonads. The first molecular evidence concerning evolutionary relationships among the fluorescent pseudomonads emerged from an immunological survey with the biotype *P. putida* as the reference strain (46). Some anomalies were associated with comparisons of CatC, but in general, serological surveys with enzymes from the β-ketoadipate pathway (36, 46) revealed relationships concordant with those discerned from nutritional analysis (45). Strains that had been assigned to *P. putida* biotype A produced enzymes that cross-reacted strongly with the reference proteins, whereas less cross-reaction was

observed with isofunctional enzymes from *P. aeruginosa*, and even less was observed with the corresponding proteins from *Pseudomonas fluorescens*. Remarkable in their absence were proteins that might have exhibited indeterminate patterns of cross-reaction. This result indicated that horizontal gene transfer had made no significant contribution to the recent evolution of genes for the β-ketoadipate pathway and raised the possibility that their genetic divergence was not entirely gradual (36).

One clearly evident genetic jump has placed *catA* (known to be several kilobase pairs away from *catC* in *P. aeruginosa* [20]) directly downstream from *catC* in *P. putida*. Conserved in both species is the capacity to express *catA* independently of the other *cat* structural genes (18, 32), and in the case of *P. putida*, this trait seems to demand superimposition of a binding site for CatR upon a DNA segment that encodes *catC*. The gene arrangement placing *catA* directly downstream from *catC* appears in *Pseudomonas* strain RB1, and by this criterion, the organism might be regarded as similar to *P. putida*. Yet nucleotide sequence divergence of 10%, accompanied by different placement of an 85-bp palindromic sequence, would argue that the genetic gap between *Pseudomonas* strain RB1 and *P. putida* warrants the generic separation accorded *E. coli* and *S. flexneri* in the enteric bacteria. Thus, the evolutionary wellspring of *Pseudomonas cat* genes sharing ancestry with counterparts in the *clc* (10) and *mdl* (51) operons is not yet defined.

Tools to sort out biological relationships among *catR,BC,A* genes of the fluorescent pseudomonads are available. PCR amplification of intergenic *cat* regions can provide direct evidence for their conservation or divergence. Suitable probes should reveal the biological distribution within the *cat* region of the 85-bp palindromic sequence that sets *Pseudomonas* strain RB1 apart from *P. putida*. Comparison of the distribution of homologs of this sequence within the DNA of related *Pseudomonas* strains may also indicate the extent to which migration of the palindromic DNA has punctuated their evolutionary divergence.

ACKNOWLEDGMENTS

J. Houghton was supported by a fellowship from the Celgene Corporation. This investigation was supported by grants from the National Science Foundation, the Army Research Office, and the General Reinsurance Corporation.

We thank Wayne Coco for thoughtful comments on the manuscript.

REFERENCES

1. **Aldrich, T. L., and A. M. Chakrabarty.** 1988. Transcriptional regulation, nucleotide sequence, and localization of the promoter of the *catBC* operon in *Pseudomonas putida*. J. Bacteriol. **170:**1297–1304.
2. **Aldrich, T. L., B. Frantz, B. Gill, J. F. Kilbane, and A. M. Chakrabarty.** 1987. Cloning and complete nucleotide sequence determination of the *catB* gene encoding *cis*,*cis*-muconate lactonizing enzyme. Gene **52:**185–195.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1988–1994. Current protocols in molecular biology J. Wiley & Sons, New York.
4. **Bagdasarian, M. M., E. Amann, R. Lurz, B. Ruckert, and M. Bagdasarian.** 1983. Activity of the hybrid *trp-lac* (*tac*) promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host range, controlled-expression vectors. Gene **26:**273–282.
5. **Berns, K. I., and C. A. Thomas, Jr.** 1965. Isolation of high molecular weight DNA from *Haemophilus influenzae*. J. Mol. Biol. **11:**476–490.
6. **Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty.** 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* cholocatechol operon in *Pseudomonas putida*. J. Bacteriol. **175:**417–427.
7. **Dimri, P. D., K. Rudd, M. K. Morgan, H. Bayat, and G. Ferro-Luzzi Ames.** 1992. Physical mapping of repetitive extragenic palindromic sequences in *Escherichia coli* and phylogenetic distribution among *Escherichia coli* strains and other bacteria. J. Bacteriol. **174:**4583–4593.
8. **Eck, R., and J. Belter.** 1993. Cloning and characterization of a gene coding for the catechol 1,2-dioxygenase of *Arthrobacter* sp. MA3. Gene **123:**87–92.
9. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
10. **Frantz, B., and A. M. Chakrabarty.** 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA **84:**4460–4464.
11. **Frantz, B., K.-L. Ngai, L. N. Ornston, and A. M. Chakrabarty.** 1987. Nucleotide sequence and expression of *clcD*, a plasmid borne diene-lactone hydrolase gene from *Pseudomonas* sp. strain B13. J. Bacteriol. **169:**704–709.
12. **Fried, M., and D. Crothers.** 1981. Equilibria kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9:**6505–6525.
13. **Gilson, E., J.-M. Clément, D. Brutlag, and M. Hofnung.** A family of dispersed repetitive extragenic palindromic DNA sequences in *E. coli*. EMBO J. **3:**1417–1421.
14. **Goldman, A., D. L. Ollis, K. Ngai, and T. A. Steitz.** 1985. Crystal structure of muconate lactonizing enzyme at 6.5 Å resolution. J. Mol. Biol. **182:**353–355.
15. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
16. **Hulton, C. S. J., C. F. Higgins, and P. M. Sharp.** 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium*, and other enteric bacteria. Mol. Microbiol. **5:**825–834.
17. **Katti, S. K., B. Katz, and H. W. Wyckoff.** 1989. Crystal structure of muconolactone isomerase at 3.3 Å resolution. J. Mol. Biol. **205:**557–571.
18. **Kemp, M. B., and G. D. Hegeman.** 1968. Genetic control of the β-ketoadipate pathway in *Pseudomonas aeruginosa*. J. Bacteriol. **96:**1488–1499.
19. **Kivissar, M., L. Kasak, and A. Nurk.** 1991. Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, pheB, of phenol degrading *Pseudomonas* sp. strain EST1001. Nature (London) **290:**754–758.
20. **Kukor, J. J., R. H. Olsen, and D. P. Ballou.** 1988. Cloning and expression of the *catA* and *catBC* gene clusters from *Pseudomonas aeruginosa* PAO. J. Bacteriol. **170:**4458–4465.
21. **Lodish, H. F.** 1970. Secondary structure of bacteriophage f2 ribonucleic acid and the initiation of *in vitro* protein biosynthesis. J. Mol. Biol. **50:**689–702.
22. **Lu, C.-D., J. E. Houghton, and A. Abdelal.** 1992. Characterization of the arginine repressor (*argR*) from *Salmonella typhimurium* and evidence for interactions with the minor groove of the *carAB* promoter. J. Mol. Biol. **225:**11–24.
23. **McFall, S. M., and M. R. Parsek.** 1994. Cross binding and cross activation by the biodegradative regulatory proteins CatR and ClcR of *Pseudomonas putida*, abstr. H-31, p. 205. *In* Abstracts of the 94th General Meeting of the American Society for Microbiology 1994. American Society for Microbiology, Washington, D.C.
24. **Meagher, R. B., and L. N. Ornston.** 1973. Relationships among the enzymes of the β-ketoadipate pathway. I. Properties of *cis*, *cis*-muconate lactonizing enzyme and muconolactone isomerase from *Pseudomonas putida*. Biochemistry **12:**3523–3530.
25. **Meighen, E. A., and R. B. Szittner.** 1992. Multiple repetitive elements and organization of the *lux* operons of luminescent terrestrial bacteria. J. Bacteriol. **174:**5371–5381.
26. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19:**269–276.
27. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Neidhart, D. J., G. L. Kenyon, J. A. Gerlt, and G. A. Petsko.** 1991. Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. Nature (London) **347:**692–694.
29. **Neidle, E. L., C. Hartnett, S. Bonitz, and L. N. Ornston.** 1988. DNA sequence of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene *catA*: evidence for evolutionary divergence of intradiol dioxygenase by acquisition of DNA repetitions. J. Bacteriol. **170:**4874–4880.
30. **Neidle, E. L., and L. N. Ornston.** 1986. Cloning and expression of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase structural gene, *catA*, in *Escherichia coli*. J. Bacteriol. **168:**815–820.
31. **Nilsson, B., M. Uhlen, S. Josephson, S. Gatenbeck, and L. Philipson.** 1983. An improved positive selection plasmid vector constructed by oligonucleotide mediated mutagenesis. Nucleic Acids. Res. **11:**8019–8030.
32. **Ornston, L. N.** 1966. The conversion of catechol and protocatechuate to β-ketoadipate by *Pseudomonas putida*. IV. Regulation. J. Biol. Chem. **241:** 3800–3810.
33. **Ornston, L. N., and D. Parke.** 1976. Properties of an inducible uptake system for β-ketoadipate in *Pseudomonas putida*. J. Bacteriol. **125:**475–488.
34. **Ornston, L. N., and R. Y. Stanier.** 1966. The conversion of catechol and protocatechuate to β-ketoadipate by *Pseudomonas putida*. I. Biochemistry. J. Biol. Chem. **241:**3776–3786.
35. **Parke, D., and L. N. Ornston.** 1976. Constitutive synthesis of enzymes of the protocatechuate pathway and of the β-ketoadipate uptake system in mutant strains of *Pseudomonas putida*. J. Bacteriol. **126:**272–281.
36. **Patel, R. N., and L. N. Ornston.** 1976. Immunological comparison of enzymes of the β-ketoadipate pathway. Arch. Microbiol. **110:**27–36.
37. **Petsko, G. A., G. L. Kenyon, J. A. Gerlt, D. Ringe, and J. W. Kozarich.** 1993. On the origin of enzymatic species. Trends Biochem. Sci. **18:**372–376.

38. **Richey, D. P., and E. C. C. Lin.** 1972. Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*. J. Bacteriol. **112:**784–790.
39. **Rothmel, R. K., T. L. Aldrich, J. E. Houghton, W. M. Coco, L. N. Ornston, and A. M. Chakrabarty.** 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the LysR family. J. Bacteriol. **172:**922–931.
40. **Rothmel, R. K., D. L. Shinabarger, M. R. Parsek, T. A. Aldrich, and A. M. Chakrabarty.** 1991. Functional analysis of the *Pseudomonas putida* regulatory protein CatR: transcriptional studies and determination of the CatR DNA-binding site by hydroxyl-radical footprinting. J. Bacteriol. **173:**4717–4724.
41. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
43. **Sharples, G. J., and R. G. Lloyd.** 1990. A novel repeated sequence located in the intergenic regions of bacterial chromosomes. Nucleic Acids Res. **18:**6503–6508.
44. **Simon, R., M. O'Connell, M. Labes, and A. Puhler.** 1986. Plasmid vectors for the genetic analysis and manipulation of rhizobia and other gram-negative bacteria. Methods Enzymol. **118:**640–659.
45. **Stanier, R. Y., N. J. Palleroni, and M. Douderoff.** 1966. The aerobic *Pseudomonads*: a taxonomic study. J. Gen. Microbiol. **48:**159–271.
46. **Stanier, R. Y., D. Wachter, C. Gasser, and A. C. Wilson.** 1970. Comparative immunological survey of two *Pseudomonas* enzymes. J. Bacteriol. **102:**351–362.
47. **Stern, M. J., G. Ferro-Luzzi Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37:**1015–1026.
48. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
49. **Tabor, S., and C. A. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
50. **Truniger, V., W. Boos, and G. Sweet.** 1992. Molecular analysis of the *glpFKX* regions of *Escherichia coli* and *Shigella flexneri*. J. Bacteriol. **174:**6981–6991.
51. **Tsou, A. Y., S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, and G. L. Kenyon.** 1990. Mandelate pathway of *Pseudomonas putida*: sequence relationships involving mandelate racemase, (S)-mandelate dehydrogenase, and benzoylformate decarboxylase and expression of benzoylformate decarboxylase in *Escherichia coli*. Biochemistry **29:**9856–9862.
52. **Versailovic, J., T. Koeuth, and J. R. Lupski.** 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of the bacterial chromosome. Nucleic Acids Res. **19:**6823–6832.
53. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13 mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
54. **Wheelis, M. L., and L. N. Ornston.** 1972. Genetic control of enzyme induction in the β-ketoadipate pathway of *Pseudomonas putida*: deletion mapping of *cat* mutations. J. Bacteriol. **108:**790–795.
55. **Wong, S. C., and A. T. Abdelal.** 1990. Unorthodox expression of an enzyme: evidence for an untranslated region within *carA* from *Pseudomonas aeruginosa*. J. Bacteriol. **172:**630–642.
56. **Wu, C. H., M. K. Ornston, and L. N. Ornston.** 1972. Genetic control of enzyme induction in the β-ketoadipate pathway of *Pseudomonas putida*. J. Bacteriol. **109:**796–802.
57. **You, I.-S., and R. Bartha.** 1982. Metabolism of 3,4-dichloraniline by *Pseudomonas putida*. J. Agric. Food Chem. **30:**274–277.
58. **Youvan, D. C., J. T. Elder, D. E. Sandlin, K. M. Szebo, D. P. Alder, N. J. Panopoulos, B. L. Marrs, and J. E. Hearst.** 1982. R-prime site directed transposon mutagenesis of the photosynthetic apparatus in *Rhodopseudomonas capsulata*. J. Mol. Biol. **162:**17–41.